

# 우리나라의 시도별 교통사고 특성분석\*

한상진\*\*

## 〈目 次〉

- |                     |                      |
|---------------------|----------------------|
| I. 서론               | III. 시도단위 사고예측모형     |
| II. 도로교통사고 특성분석 방법론 | 1. 시계열분석 모형          |
| 1. 사고예측모형의 종류 및 내용  | 2. 회귀분석 모형 (횡단면 분석)  |
| 2. 시계열분석 모형         | 3. 인구규모별 도로교통사고 특성분석 |
| 3. 회귀분석 모형          | IV. 결론               |

## I. 서론

2006년 경찰청의 도로교통사고 통계에 의하면 전체 교통사고 213,745건 중 71%에 달하는 151,596건이 특별·광역시도와 시·군도에서 발생하였다. 이는 중앙정부에서 관리하는 고속도로나 일반국도보다 지방자치단체에서 관리하는 도로에서 상대적으로 많은 사고가 발생함을 의미하며 그만큼 교통사고감소와 예방을 위한 지방자치단체의 노력이 중요함을 보여준다.

본 연구에서는 특별시, 광역시, 도 등 광역자치단체의 도로교통사고 특성을 설명하는 거시적 차원의 사고예측모형을 개발하여 도로교통사고 감소방안 마련의 방향을 제시하고자 한다. 거시적 사고예측모형은 크게 횡단면 자료에 기반을 둔 회귀분석모형과 시계열 자료에 기반을 둔 시계열분석 모형으로 나누어 제시한다. 회귀분석모형은 인구, 자동차등록대수 등 거시적 사회·경제변수가 교통사고발생에 어느 정도 영향을 미치는지를 분석하여 교통안전정책의 방향을 찾기 위해 실시되며 시계열분석모형은 특정 시·도의 교통사고발생 추이를 파악하여 현재의 교통사고 추세가 중장기적으로 어떻게 변화할지를 추정하기 위해 실시된다. 한편, 시·도별 인구규모의 차이가 교통사고발생에 미치는 영향도 함께 분석하여 본다.

\* 본 논문은 건설교통부가 출연하고 한국건설교통기술평가원에서 위탁시행 한 2005년도 건설기술기반구축사업(05기반구축D02)의 지원으로 이루어졌습니다.

\*\* 한국교통연구원

## II. 도로교통사고 특성분석 방법론

### 1. 사고예측모형의 종류 및 내용

사고예측모형은 크게 거시적 모형과 미시적 모형으로 구분된다. 거시적 사고예측모형은 국가, 지방자치단체, 도로관리청 등 특정한 지리적 권역에서 발생하는 도로교통사고의 발생추이를 예측하기 위해 구축된다. 거시적 모형에 주로 이용되는 자료는 도로교통사고 발생건수, 사망자수, 부상자수 등의 연도별, 월별, 주별 등 시계열 자료이다. 이 밖에 인구, 자동차등록대수, 지역총생산, 도로연장 등 주요 사회·경제지표도 활용된다. 시계열 자료에 대한 분석은 단순 회귀분석법이나 자기회귀모형(Autoregression, AR), 이동평균법(Moving Average, MA), 이 둘을 결합시킨 자기이동평균모형(ARMA), ARMA 모형에 차분과정을 거치는 ARIMA 모형 등을 활용한다. 거시적 사고예측모형은 장래 교통사고의 발생추이를 추정하여 교통사고 감소목표를 설정할 때 활용되거나 교통사고감소를 위해 주로 관리해야 할 정책변수를 선별할 때에도 활용된다.

미시적 사고예측모형은 주로 도로교통사고와 도로환경적 요소와의 관계를 규명하기 위해 구축된다. 주로 사용되는 변수는 교통량(연평균일교통량, AADT), 중차량구성비 등 교통관련변수, 설계속도, 평면곡선길이, 편구배, 종단곡선길이, 종단경사, 길어깨폭, 차로폭 등 도로선형관련 변

〈표 1〉 거시적 사고예측모형과 미시적 사고예측모형의 비교

| 구분       | 거시적 사고예측모형   | 미시적 사고예측모형   |
|----------|--|--|
| 모형의 목적   | <ul style="list-style-type: none"> <li>· 국가, 지방정부, 도로관리청 차원의 도로교통 사고 발생 경향 파악</li> <li>· 인구, 자동차등록대수 등 사회경제지표와 교통사고 발생의 관계분석을 통한 주요 정책 변수의 파악</li> </ul> | <ul style="list-style-type: none"> <li>· 도로교통사고와 도로의 선형, 운영방식, 교통량 등 도로환경과의 관계를 규명</li> </ul>  |
| 자료의 종류   | <ul style="list-style-type: none"> <li>· 사고건수 추이(연, 월, 일 단위)</li> <li>· 사회·경제지표(인구, 자동차등록대수, GDP, 대중교통수단분담율 등)</li> </ul>                                | <ul style="list-style-type: none"> <li>· 교통관련요소: 교통량(AADT), 중차량구성비 등</li> <li>· 도로선형요소: 평면곡선, 편구배, 종단곡선길이, 종단경사, 길어깨 폭, 차로폭, 설계속도 등</li> <li>· 도로운영요소: 신호주기, 현시 등</li> </ul> |
| 분석방법     | <ul style="list-style-type: none"> <li>· 시계열분석: AR, MA, ARMA, ARIMA 모형</li> <li>· 선형/비선형 회귀분석</li> </ul>   | <ul style="list-style-type: none"> <li>· 선형/비선형 회귀분석</li> <li>· 포아송분포</li> <li>· 음이항분포</li> </ul>  |
| 모형의 활용방안 | <ul style="list-style-type: none"> <li>· 교통사고의 장래발생 추이 추정 및 교통사고 감소 목표 설정(시계열 자료)</li> <li>· 교통사고 감소를 위한 주요정책변수 선정(횡단면 자료)</li> </ul>                    | <ul style="list-style-type: none"> <li>· 도로계획 단계에서 노선대안별 안전도 평가</li> <li>· 도로설계 단계에서 도로선형 등 설계요소 변화가 사고에 미치는 영향파악</li> </ul>   |

수, 신호주기길이, 유효녹색시간길이 등 도로운영관련 변수 등이 있다. 이러한 변수와 교통사고와의 관계는 선형 혹은 비선형 회귀분석모형을 통해 구축된다. 이 때 교통사고의 발생분포는 대개 포아송 혹은 음이항분포로 가정한다. 미시적 사고예측모형은 도로계획 단계에서 노선대안별 안전도를 평가하는데 활용하거나 도로설계단계에서 일부 설계요소의 변화(가령, 평면곡선의 길이 축소)가 사고에 미치는 효과를 추정하기 위해 사용된다. <표 1>은 거시적 사고예측모형과 미시적 사고예측모형을 비교설명하고 있다.

## 2. 시계열분석 모형

국민총생산, 물가지수, 주가지수 등과 같이 시간에 따라 지표가 산출되는 자료를 시계열 자료라 부르며 이러한 자료에 대한 분석기법을 시계열 분석기법이라 부른다. 이 기법은 Box와 Jenkins(1976)에 의한 ARIMA 모형 이후 비약적인 발전을 이루고 있다.

시계열 모형 중 대표적인 모형으로는 AR(자기회귀, Auto-Regressive)모형과 MA(이동평균법, Moving Average)모형 그리고 이들을 결합시킨 ARMA 모형이 있다. 또한 원시 시계열자료를 정상시계열 자료로 전환시키는 차분과정을 포함하는 ARIMA 모형이 있다.

평균이  $\mu$ 인 AR(1), MA(1), ARMA(1, 1)모형은 다음과 같이 표현한다.

$$\cdot \text{평균이 } \mu \text{인 AR(1)} : (Z_t - \mu) = \phi(Z_{t-1} - \mu) + a_t \quad (1a)$$

$$\cdot \text{평균이 } \mu \text{인 MA(1)} : Z_t - \mu = a_t - \theta a_{t-1} \quad (1b)$$

$$\cdot \text{평균이 } \mu \text{인 ARMA(1, 1)} : Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t - \theta a_{t-1} \quad (1c)$$

일반적으로 시계열 분석 모형을 수립하기 위해 3단계인 모형식별, 모수추정, 모형진단 과정을 거친다. 1단계인 모형의 식별단계에서는 표본자기상관계수(SACF)가 급격히 0에 가까워지는 절단점으로 AR(p)모형을 선택 할 수 있고, MA(q)모형은 표본부분자기상관계수(SPACF)가 지수적으로 급격히 0에 가까워지는 절단점으로 선정할 수 있다. 만약 비정상시계열인 경우에는 정상시계열로 전환시키기 위해 차분과정(d)을 거치며 이 경우 ARIMA(p, d, q) 모형을 선택한다. 일반적으로 시계열 모형은 여러 가지 대안으로 요약될 수 있는데 이중 가장 적합한 모형을 선택하기 위해서는 대개 AIC(Akaike Information Criteria; Akaike 정보량 기준)값이 최소가 되는 모형을 선택한다.

2단계는 모수추정단계로 조건부최소제곱추정법(Conditional Least Squares; CLS)을 이용하여 추정한다. 이렇게 결정된 예측모형의 모수추정치가 유의한지를 판단하기 위해 t-검정을 실시한다. t-검정은 귀무가설( $H_0: \theta \text{ or } \phi = 0$ )로 추정된 모수가 0일 가능성이 높은지의 여부를 판단한다.

3단계 모형진단단계에서는 모형의 과다적합진단(overfitting diagnostics)과 잔차분석(residual analysis)을 통해 모형의 적합성을 최종 결정한다(조신섭 · 황선영, 1997).

### 3. 회귀분석 모형

회귀분석 모형은 가장 일반적으로 활용되는 기법으로 시계열자료(time-series data)를 이용하여 교통사고자료의 추세를 설명하는데 활용될 수도 있고 횡단면자료 (cross-sectional data)를 이용하여 교통사고자료와 사회·경제지표와의 관계를 설명할 수 있는 모형을 구축할 수도 있다. 시간  $t$ 에 대한 단순선형회귀모형(simple linear regression model)은 식 (1)과 같이 나타낼 수 있다. 횡단면자료를 이용할 때에는 독립변수  $t$ 가 인구, 자동차등록대수 등 사회·경제변수로 바뀌면 된다.

$$y_i = \beta_0 + \beta_1 t + \varepsilon_i \quad (2)$$

여기서,  $y_i = i$  연도의 사고건수(비율)

$\beta_0$  = 직선의 절편

$\beta_1$  = 직선의 기울기

$\varepsilon_i = i$  연도 값의 오차

## III. 시도단위 사고예측모형

### 1. 시계열분석 모형

본 연구에서는 통계분석 소프트웨어인 SAS를 이용하여 인천광역시, 부산광역시, 광주광역시, 대구광역시에서 1989년부터 2004년까지 도로교통사고 사망자수 추이를 시계열분석 모형 (ARIMA)으로 추정하였다. 그 결과 모든 추세가 1차 자기회귀분석모형(AR(1))으로 설명되었다. <표 2>~<표 5>는 이러한 결과를 요약하고 있다.

이들 표에는 모형의 모수추정치( $\hat{\phi}_1, \hat{\mu}$ ), 이들의 유의성을 설명하는 통계량(t-value, p-value), 그리고 모형의 적합도를 설명하는 AIC가 정리되었으며, 모형식과 향후 5년간의 예측치도 제시되어 있다. 인천광역시와 부산광역시의 경우는 자기상관계수의  $t$ 값이 다소 낮고, 유의확률이 높게 나타났으나 대체로 모형의 설명력은 양호한 것으로 판단된다.

<표 2> 인천광역시 도로교통사고 추세모형

| 구 분            |                  | estimate                                 | standard error | T value               | Pr >  T  | AIC      |
|----------------|------------------|--|----------------|-----------------------|----------|----------|
| 모수추정치<br>& 통계량 | $\hat{\mu}$      | 238.53535                                | 26.27527       | 9.08                  | <0.0001  | 156.1364 |
|                | $\hat{\sigma}_1$ | 0.66782                                  | 0.22524        | 2.96                  | 0.0110   |          |
| 모형식            |                  | $Z_t = 238.53535 + 0.66782Z_{t-1} + a_t$ |                |                       |          |          |
| 예측치<br>(5년)    | 연도               | Forecast (명)                             | Std Error      | 95% Confidence Limits |          |          |
|                | 2005             | 202                                      | 41.4225        | 120.2612              | 282.6343 |          |
|                | 2006             | 214                                      | 49.8101        | 116.1415              | 311.3936 |          |
|                | 2007             | 222                                      | 56.1256        | 117.8707              | 326.1191 |          |
|                | 2008             | 228                                      | 54.5393        | 120.5943              | 334.3843 |          |
|                | 2009             | 232                                      | 55.1581        | 123.0508              | 339.2664 |          |

〈표 3〉 부산광역시 도로교통사고 추세모형

| 구 분            |                  | estimate                                 | standard error | T value               | Pr >  T | AIC      |
|----------------|------------------|--|----------------|-----------------------|---------|----------|
| 모수추정치<br>& 통계량 | $\hat{\mu}$      | 388.5505                                 | 41.21844       | 9.43                  | <0.0001 | 170.1142 |
|                | $\hat{\sigma}_1$ | 0.66267                                  | 0.24456        | 2.71                  | 0.0179  |          |
| 모형식            |                  | $Z_t = 388.55050 + 0.66267Z_{t-1} + a_t$ |                |                       |         |          |
| 예측치<br>(5년)    | 연도               | Forecast (명)                             | Std Error      | 95% Confidence Limits |         |          |
|                | 2005             | 306.6770                                 | 66.0064        | 177.3068              |         | 436.0472 |
|                | 2006             | 334.2952                                 | 79.1839        | 179.0976              |         | 489.4928 |
|                | 2007             | 352.5970                                 | 84.3224        | 187.3282              |         | 517.8658 |
|                | 2008             | 364.7251                                 | 86.4824        | 195.2227              |         | 534.2276 |
|                | 2009             | 372.7621                                 | 87.4141        | 201.4336              |         | 544.0906 |

〈표 4〉 광주광역시 도로교통사고 추세모형

| 구 분            |             | estimate                                 | standard error | T value               | Pr >  T | AIC      |
|----------------|-------------|--|----------------|-----------------------|---------|----------|
| 모수추정치<br>& 통계량 | $\hat{\mu}$ | 194.61938                                | 23.90750       | 8.14                  | <.0001  | 147.642  |
|                | $\Phi_1$    | 0.77219                                  | 0.20668        | 3.74                  | 0.0025  |          |
| 모형식            |             | $Z_t = 194.61939 + 0.77219Z_{t-1} + a_t$ |                |                       |         |          |
| 예측치<br>(5년)    | 연도          | Forecast (명)                             | Std Error      | 95% Confidence Limits |         |          |
|                | 2005        | 150.1261                                 | 31.2081        | 88.9593               |         | 211.2929 |
|                | 2006        | 160.2619                                 | 39.4296        | 82.9813               |         | 237.5426 |
|                | 2007        | 168.0888                                 | 43.6003        | 82.6338               |         | 253.5438 |
|                | 2008        | 174.1326                                 | 45.9072        | 84.1561               |         | 264.1091 |
|                | 2009        | 178.7996                                 | 47.2292        | 86.2321               |         | 271.3672 |

〈표 5〉 대구광역시 도로교통사고 추세모형

| 구 분            |             | estimate                                 | standard error | T value               | Pr >  T | AIC      |
|----------------|-------------|--|----------------|-----------------------|---------|----------|
| 모수추정치<br>& 통계량 | $\hat{\mu}$ | 399.51251                                | 58.06674       | 6.88                  | <.000.1 | 166.5588 |
|                | $\Phi_1$    | 0.85107                                  | 0.15463        | 6.15                  | <.000.1 |          |
| 모형식            |             | $Z_t = 399.51251 + 0.85107Z_{t-1} + a_t$ |                |                       |         |          |
| 예측치<br>(5년)    | 연도          | Forecast (명)                             | Std Error      | 95% Confidence Limits |         |          |
|                | 2005        | 232.5879                                 | 58.6296        | 117.6760              |         | 347.4998 |
|                | 2006        | 240.7556                                 | 80.9117        | 82.1717               |         | 399.3396 |
|                | 2007        | 248.5237                                 | 96.7426        | 58.9116               |         | 438.1357 |
|                | 2008        | 255.9117                                 | 109.1012       | 42.0772               |         | 469.7461 |
|                | 2009        | 262.9381                                 | 119.1812       | 29.3473               |         | 496.5290 |

〈표 2〉~〈표 5〉에 제시된 AR(1) 모형을 활용하면 향후 5년 이외에도 장래의 도로교통사고 사망자수 추정하는데 도움이 될 것이다. 그러나 시간이 늘어날수록 예측의 정확도는 떨어지는 경향이 발생하며 AR(p) 모형은 장기적으로 평균에 수렴하는 경향이 있어 지나치게 긴 시차에 대한 예측은 바람직하지 않음에 유의하여야 한다.

## 2. 회귀분석 모형(횡단면 분석)

본 연구에서는 주요 사회·경제지표와 교통사고 사망자수의 관계식을 도출하기 위해 2004년을 기준으로 16개 광역시·도에 대한 국토면적당 도로연장, 면적당 유효도로연장, 인구 천명당 도로연장, 인구당 자동차보유대수, 인구당 대중교통차량대수, 도로교통사고 사망자 수를 〈표 6〉과 같이 정리하였다. 그러나 도시의 소득규모를 나타내는 지역총생산(GRP) 지표를 구하지 못해 소득과 교통사고와의 관계는 규명하지 못하였다.

〈표 6〉 주요 사회·경제지표와 도로교통사고 사망자수

| 지역 | 국토면적당<br>도로연장(km/km <sup>2</sup> ) | 지역별 면적당<br>유효도로연장 | 인구 천명당<br>도로연장 | 인구당 자동차<br>보유대수 | 인구당<br>대중교통<br>차량대수 | 교통사고<br>사망자수 |
|----|------------------------------------|-------------------|----------------|-----------------|---------------------|--------------|
| 서울 | 13.232                             | 12.3              | 0.78           | 270.2           | 7.91                | 468          |
| 부산 | 3.855                              | 5.16              | 0.8            | 262.3           | 7.70                | 265          |
| 대구 | 2.477                              | 4.54              | 0.86           | 327.5           | 7.33                | 224          |
| 인천 | 2.216                              | 4.17              | 0.84           | 299.9           | 5.78                | 183          |
| 광주 | 2.490                              | 5.49              | 0.89           | 300.2           | 6.50                | 137          |
| 대전 | 2.994                              | 4.73              | 1.11           | 339             | 6.62                | 98           |
| 울산 | 1.363                              | 2.26              | 1.32           | 340             | 5.62                | 95           |
| 경기 | 1.250                              | 2.36              | 1.19           | 316             | 4.01                | 1166         |
| 강원 | 0.554                              | 0.64              | 6.11           | 336.1           | 5.95                | 224          |
| 충북 | 0.880                              | 1.44              | 4.36           | 327.9           | 5.47                | 296          |
| 충남 | 0.877                              | 1.31              | 3.82           | 322.4           | 3.94                | 537          |
| 전북 | 0.915                              | 1.54              | 3.84           | 303.1           | 5.57                | 395          |
| 전남 | 0.838                              | 1.09              | 5.06           | 285.4           | 4.77                | 579          |
| 경북 | 0.621                              | 0.85              | 4.35           | 336.6           | 4.49                | 698          |
| 경남 | 1.147                              | 1.25              | 3.81           | 321.2           | 5.08                | 494          |
| 제주 | 1.731                              | 1.68              | 5.74           | 370.3           | 10.41               | 91           |

자료: 설재훈 외, 『도로교통 부문의 국가경쟁력 강화방안: 국제 및 지역간 비교를 중심으로』, 한국교통연구원, 2005.

〈표 6〉에서 제시한 사회·경제지표와 도로교통사고 사망자수의 관계를 모형화하기 위해서는 우선 독립변수로 사용될 사회·경제지표들 간의 상관도를 분석할 필요가 있다. 상관도가 높은 모형이 동시에 사용되면 다중공선성 때문에 모형의 설명력과 객관성이 떨어지기 때문이다. 〈표 7〉은 6개 지표간의 피어슨(Pearson) 상관계수를 보여주고 있다. 피어슨 상관계수  $R$ 은  $-1 \leq R \leq 1$ 의 값을 갖는다.  $-1$  혹은  $1$ 에 가까울수록 상관도가 높고  $0$ 에 가까울수록 상관도가 낮다고 판단할 수 있다. 상관도 분석결과 국토면적당도로연장과 지역별면적당유효도로연장 사이에는 높은 상관계수( $0.95$ )가 도출되어 국토면적당 도로연장만 모형에 사용하기로 하였다.

선정된 4개 사회·경제지표로 일반 선형회귀분석 모형으로 모수를 추정한 결과는 〈표 8〉과 같다. 그러나 이 경우 인구당 대중교통차량대수를 제외한 모든 변수의  $t$ 값과  $R^2$ 값 ( $0.45$ )이 낮아 설명력이 매우 떨어진다고 판단된다. 또한 인구당자동차보유대수의 모수가 음수가 나와 모형의 현실성도 떨어진다.

〈표 7〉 선정된 사회·경제지표 사이의 피어슨 상관계수

| 지역                              | 국토면적당<br>도로연장<br>(km/km <sup>2</sup> ) | 지역별 면적당<br>유효도로연장 | 인구<br>천명당<br>도로연장 | 인구당<br>자동차<br>보유대수 | 인구당<br>대중교통<br>차량대수 | 사망자수     |
|---------------------------------|--|-------------------|-------------------|--------------------|---------------------|----------|
| 국토면적당 도로연장(km/km <sup>2</sup> ) | 1.00000                                | 0.94681           | -0.47549          | -0.51334           | 0.44371             | -0.05116 |
| 지역별 면적당 유효도로연장                  | 0.94681                                | 1.00000           | -0.68808          | -0.55165           | 0.45108             | -0.15348 |
| 인구 천명당 도로연장                     | -0.47549                               | -0.68808          | 1.00000           | 0.44165            | -0.08558            | 0.06758  |
| 인구당 자동차 보유대수                    | -0.51334                               | -0.55165          | 0.44165           | 1.00000            | 0.09023             | -0.19961 |
| 인구당 대중교통 차량대수                   | 0.44371                                | 0.45108           | -0.08558          | 0.09023            | 1.00000             | -0.60579 |
| 사망자수                            | -0.05116                               | -0.15348          | 0.06758           | -0.19961           | -0.60579            | 1.00000  |

〈표 8〉 모수추정치 및 통계량(다중 선형회귀분석)

| Variable                      | Paramter<br>Estimate | Standard Error | T value | Pr >  T | R-Square | Adj R-Sq |
|-------------------------------|----------------------|----------------|---------|---------|----------|----------|
| Intercept                     | 1105.88766           | 882.64922      | 1.25    | 0.2362  | 0.4507   | 0.2509   |
| 면적당 도로연장(km/km <sup>2</sup> ) | 32.27272             | 31.01217       | 1.04    | 0.3204  |          |          |
| 시도별 인구 천명당 도로연장               | 26.42406             | 37.61152       | 0.70    | 0.4969  |          |          |
| 인구당 자동차 보유대수                  | -0.37177             | 2.97047        | -0.13   | 0.9027  |          |          |
| 인구당 대중교통 차량대수                 | -126.17576           | 46.74900       | -2.70   | 0.0207  |          |          |

〈표 9〉 모수추정치 및 통계량(원점강제 회귀분석)

| 변수                            | 모수추정치      | 표준오차     | t value | Pr >  t | R <sup>2</sup> | Adj R <sup>2</sup> |
|-------------------------------|------------|----------|---------|---------|----------------|--------------------|
| 면적당 도로연장(km/km <sup>2</sup> ) | 50.99029   | 27.81450 | 1.83    | 0.0917  | 0.7776         | 0.7034             |
| 시도별 인구 천명당 도로연장               | 21.41127   | 38.27572 | 0.56    | 0.5862  |                |                    |
| 인구당 자동차 보유대수                  | 3.18054    | 0.90691  | 3.51    | 0.0043  |                |                    |
| 인구당 대중교통 차량대수                 | -134.77708 | 47.32734 | -2.85   | 0.0147  |                |                    |

이러한 문제를 극복하기 위해 원점강제 회귀분석을 실시하였다. 〈표 9〉는 원점강제 회귀분석의 모수추정치 및 통계량을 보여주고 있다. 특히 도로교통사고 사망자수가 음수인 경우는 없으므로 이러한 접근법은 타당하다고 볼 수 있다. 이 경우에는  $R^2$ 값도 0.78로 우수한 편이고, 모수의  $t$ 값도 인구천명당도로연장을 제외하고는 대체로 양호한 것으로 판단된다. 이렇게 도출된 도로교통사고 사망자수(Y)와 사회·경제지표의 관계는 식 (3)과 같이 표현된다. 식 (3)에서 볼 수 있듯 도로교통사고 사망자 수를 증가시키는 주요 지표는 면적당도로연장이며, 다음으로 인구천명당도로연장 및 인구당자동차보유대수인 것으로 나타났다. 반면, 인구당대중교통차량대수가 높을수록 사망자수는 감소하는 경향이 있음을 알 수 있다.

이러한 결과는 도로교통사고 사망자수를 줄이기 위해서는 승용차 중심의 도로교통정책에 치중하기보다 대중교통중심의 정책에 더 큰 관심을 기울여야 함을 함의한다고 볼 수 있다.

$$Y = 50.99029(\text{면적당도로연장}) + 21.41127(\text{시도별 인구천명당도로연장}) + 3.18054(\text{인구당자동차보유대수}) - 134.77708(\text{인구당 대중교통차량대수}) \quad (3)$$

### 3. 인구규모별 도로교통사고 특성분석

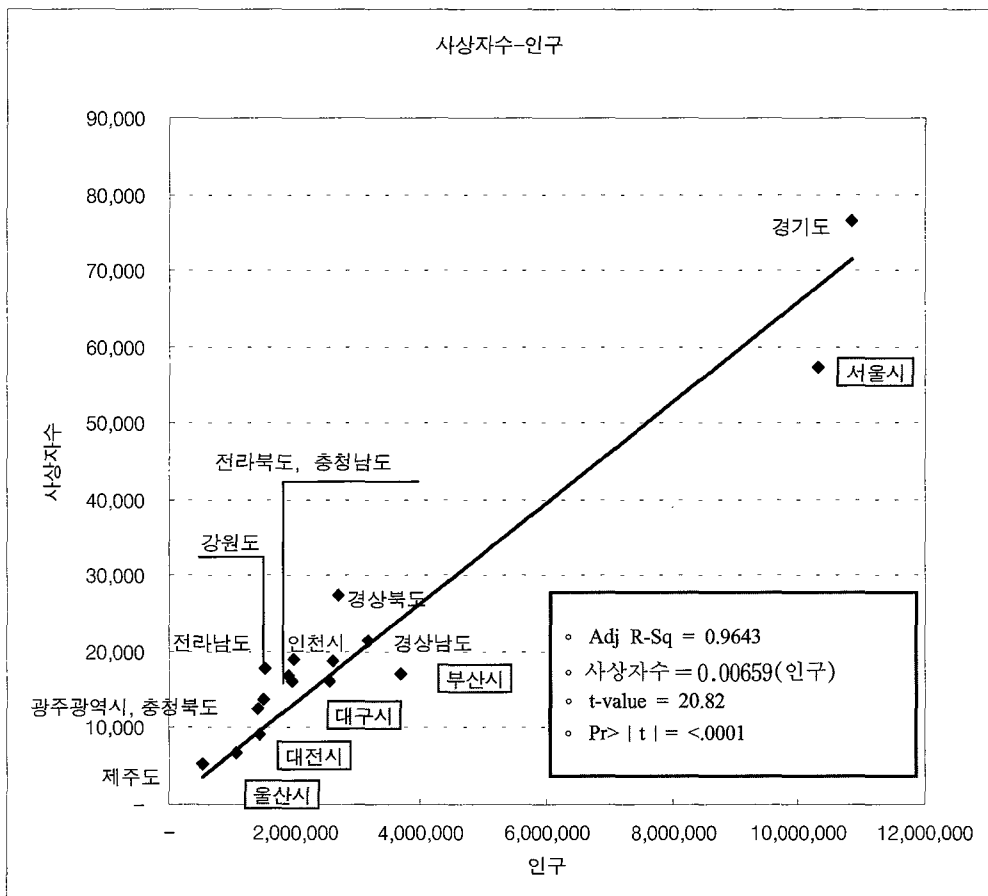
도로교통사고는 대개 인구가 많을수록, 차량등록대수가 높을수록, 도로연장이 길수록 많이 발생하는 경향이 강하다. 따라서 지역별로 교통사고의 위험도를 비교할 때는 인구당 사망자수, 차량당 사망자수, 혹은 도로 km당 사망자수 등이 많이 활용된다. 이렇게 하면 단순전수의 비교가 갖는 한계를 극복할 수 있어 국가간 교통사고의 위험도를 비교할 때에도 활용된다. 그러나 한 나라에서 인구규모별 교통사고의 위험도를 비교할 때는 회귀분석모형을 활용할 수도 있다.

〈그림 1〉은 2004년 시·도별 인구와 교통사고 사상자수의 관계를 나타내고 있다. 그림에서 직선식은 교통사고 사상자수와 도시인구의 관계를 설명하는 회귀식이다. 이 회귀식의 모형 적합도는  $R^2$ 값이 0.96으로 매우 높고, 모수의  $t$ 값도 높은 것으로 나타나 통계적으로 유의하다고 판단할 수 있다.

이 회귀식과 가까운 곳에 위치한 시·도는 비교적 인구규모에 맞는 수준의 교통사고 사상자



가 발생하고 있다고 판단할 수 있지만 회귀식보다 위에 위치한 시·도의 경우는 인구규모에 비해 교통사고 사상자수가 높다고 판단할 수 있다. 반대로 회귀식보다 아래에 위치한 시·도의 경우는 비교적 안전하다고 판단할 수 있다. 인구에 비해 도로교통사고 사상자수가 많은 시·도에는 경상북도, 강원도, 전라남도, 경기도, 인천광역시 등이 있다. 부산광역시, 서울특별시 등은 인구규모에 비해 비교적 도로교통사고 사상자 수가 적게 나타난 것으로 분석되었다. 이런 식의 분석은 인구규모별 사고 위험도를 직선에서 얼마나 떨어져 있는지를 통해 쉽게 파악할 수 있어 유용하다. 특히 직선과의 차이만큼을 도로교통사고 사상자수 감소 목표치(target)로 정할 수도 있다. 이런 식의 분석은 인구규모 이외에도 자동차등록대수, 도로총연장에 대해서도 시행할 수 있다.



〈그림 1〉 교통사고 사상자수와 도시인구와의 관계(시·도별 비교)

#### IV. 결론

본 연구에서는 사고예측모형의 종류를 거시적 차원과 미시적 차원으로 나누어 설명한 후 시계열분석 모형과 회귀분석 모형으로 우리나라 시·도별 도로교통사고 사망자수 예측모형을 제시하였다. 그 결과 시계열분석 모형은 대체로 AR(1) 모형으로 적합되었다. 횡단면 자료를 통해 구축한 회귀분석모형은 시·도의 차량대수 및 도로연장이 높을수록 도로교통사고 사망자는 늘어나고 반대로 대중교통차량대수가 높은 경우에는 도로교통사고 사망자수가 줄어들 수 있음을 보여주었다. 또한, 회귀분석 모형을 통해 인구규모별로 적절한 수준의 도로교통사고 사상자수는 어느 수준인지도 평가하여 상대적으로 안전한 시·도를 파악하여보았다.

그러나 연단위 자료에 기반한 시계열분석 모형은 활용성이 낮을 것으로 보인다. 우선 최종적으로 선정된 시계열 분석모형이 정상시계열 모형인 AR 모형으로 적합되었기 때문에 장래 추정치가 시간이 흐를수록 평균치로 회귀할 가능성이 높다. 이는 모형의 실효성이 떨어짐을 의미한다. 실제로 본 연구에서 제시한 장래 5년 동안의 예측사고건수는 지속적으로 늘어나는 경향을 띄고 있다. 이러한 문제를 완화하기 위해서는 월단위 자료를 통해 자료의 수를 늘리는 방안을 검토할 수 있다. 횡단면 자료를 이용한 회귀분석 모형은 지역총생산(GRP)과 같은 소득관련 지표를 추가적으로 검토할 필요가 있다. 또한 인구규모 이외의 차량등록대수, 도로총연장별 교통사고 위험수준의 비교도 추가적으로 연구될 필요가 있을 것이다.

#### 참고문헌

- 도로교통안전관리공단, 교통사고통계분석(1990~2005).
- 조신섭, 황선영, 시계열 분석, 2003.
- 송문섭 외 3명, SAS를 이용한 통계자료분석, 1992.
- 도로교통안전관리공단, 1990년~2005년 교통사고 통계분석, 1990~2005.
- 권혁제, SAS 활용 표준통계학, 2000.
- 한상진, 김근정, “도로종류별 교통사고 추세분석 및 시계열 분석모형 개발”, 한국도로학회, 제9권 제3호, 2007.
- Box, G.E.P. and G.M. Jenkins (1976) Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco: Holden-Day.